Using Vision for Animating Virtual Humans

Ioannis Kakadiaris

Visual Computing Lab
Department of Computer Science
University of Houston

1

Visual Computing

- The field of Visual Computing is concerned with the analysis, numerical manipulation, querying, display, storage, and transmission of data.

2

Application Domains

- Human Motion Analysis
- Biomedical Data Analysis
- Seismic Data Analysis

3

Overview

- ✓ Motivation
- Theoretical Framework
  - Distributed Approximating Functionals
  - Physics-Based Modeling
- Human Motion Analysis
- Biomedical Data Analysis
- Conclusion

4

Physics-Based Models: Computer Vision

- Objective
  - Represent nonrigid shapes
  - Reconstruct nonrigid shapes from noisy data
  - Estimate the motion of nonrigid objects
- Solution
  - Use the principles of physics to approximate the shape of objects and their behavior

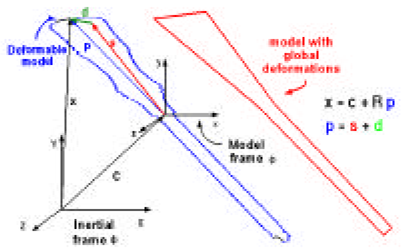
5

Physics-Based Models: Computer Graphics

- Objective
  - Model nonrigid objects and their interaction with the physical world
  - Realistically simulate and animate the motion of articulated objects with deformable parts
- Previous Attempts
  - Geometric modeling techniques have had limited success
- Solution
  - A mathematical representation of an object (or its behavior) which incorporates physical characteristics such as forces, torques and energies into the model allowing numerical simulation of its behavior.

6

Geometry of Rigid/Deformable Models



Geometry: Global Deformations

- Geometric primitive: $\mathbf{e}(\mathbf{u}; a_1, a_2, \ldots)$
- Parameterized deformations:

$$\mathbf{T}(\mathbf{e}; b_1, b_2, \ldots)$$

- Global deformation parameter vector

$$\mathbf{s} = \mathbf{T}(\mathbf{e}(\mathbf{u}; a_1, a_2, \ldots); b_1, b_2, \ldots)$$

$$\mathbf{q}_s = (a_1, a_2, \ldots, b_1, b_2, \ldots)^T$$

Geometry: Global Deformations (cont.)

- Example: Superquadric

Geometry: Global Deformations (cont.)

- Example: Superquadric with deformation

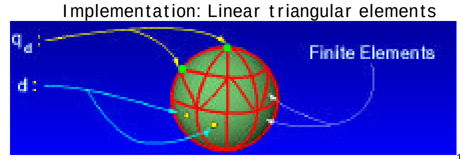
Geometry: Local Deformations

- Finite elements
  - Local deformation: **d**
  - Linear combination of nodal displacements:

$$\mathbf{d} = \mathbf{S}\mathbf{q}_d$$

S: Matrix of local finite element shape functions

Implementation: Linear triangular elements



Kinematics

- Generalized coordinate vector

$$\mathbf{q} = (\mathbf{q}_c^T, \mathbf{q}_\theta^T, \mathbf{q}_s^T, \mathbf{q}_d^T)^T$$

translation rotation global-def local-def

- Velocity of points on the model

$$\mathbf{x} = [\mathbf{I}\ \mathbf{B}\ \mathbf{RJ}\ \mathbf{RS}]\mathbf{q} = \mathbf{Lq}$$

- Jacobian $\mathbf{L}$ maps from $\mathbf{q}$-space to 3-space

Dynamics

- Lagrange equations of motion

Vision-Shape: $\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{f}_q$

Vision-Motion: $\ddot{\mathbf{q}} + \dot{\mathbf{q}} = \mathbf{f}_q$

Graphics: $\mathbf{M}\ddot{\mathbf{q}} + \mathbf{D}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{f}_q + \mathbf{g}_q$

$\mathbf{M}$: block symmetric mass matrix
$\mathbf{D}$: Raleigh damping matrix, $\mathbf{D} = a\mathbf{M} + b\mathbf{K}$
$\mathbf{K}$: stiffness matrix
$\mathbf{f}_q(\mathbf{u},t)$: generalized external forces
$\mathbf{g}_q(\mathbf{u},t)$: generalized inertial forces

13

Dynamics: Generalized Forces

- Generalized external forces

$$\mathbf{f}_q = \mathbf{L}^T\mathbf{f} = \left(\underbrace{\mathbf{f}_c^{\,T}}_{\text{translation-f}}, \underbrace{\mathbf{f}_q^{\,T}}_{\text{rotation-f}}, \underbrace{\mathbf{f}_s^{\,T}}_{\text{global-f}}, \underbrace{\mathbf{f}_d^{\,T}}_{\text{local-f}}\right)^T$$

14

Dynamics: Generalized Forces (cont.)

- Generalized inertial forces

$$\mathbf{g}_q = -\int \mu\, \mathbf{L}^T \dot{\mathbf{L}}\, \dot{\mathbf{q}}\, d\mathbf{u}$$

where

$$\dot{\mathbf{L}}\dot{q} = w \times (w \times \mathrm{Rp}) + 2w \times \mathrm{R}\dot{\mathrm{p}}$$

15

Numerical Simulation of Motion Equations

- Second order system

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{D}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{g}_q + \mathbf{f}_q + \mathbf{f}_{gc}$$

- Numerically integrate through time

16

Overview

- ✓ Motivation
- ✓ Theoretical Framework
- Human Motion Analysis
  - Inferring Structure in 2D
  - Human Body Model Acquisition
  - Human Body Tracking
  - Estimating Anthropometry and Pose from a Single Camera
- Biomedical Data Analysis
- Conclusion

17

Motion-Based Part Segmentation

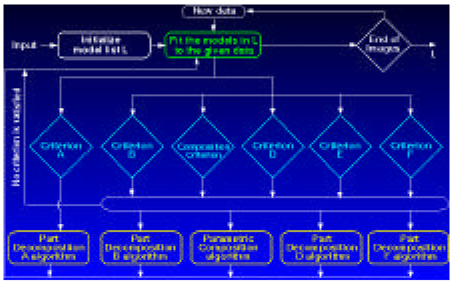
- **Given an image sequence of a multi-part object whose parts move relative to one another …**

  Recover a structured description in terms of moving parts, without a priori knowledge of the object or the object domain.

  Accurately estimate the parts' shape and motion parameters.

18

Part Segmentation Algorithm (PSA)



19

Motion-Based Part Segmentation

Advantages

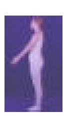
- Integrates the processes of part segmentation and fitting
- Allows reliable shape description of the parts
- Estimates the location of the joints between the parts (if any)
- Detects multiple joints
- Does not require an a priori model of the multi-part object or of the shape of the parts

20

Contribution

- New framework for the two-dimensional part segmentation shape and motion estimation of multi-part objects.





21

Overview

- ✓ Motivation
- ✓ Theoretical Framework
- Human Motion Analysis
  - Inferring Structure in 2D
  - Human Body Model Acquisition
  - Human Body Tracking
  - Estimating Anthropometry and Pose from a Single Camera
- Biomedical Data Analysis
- Conclusion

22

Human Model Acquisition

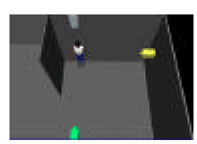
Given image sequences (from multiple views) of a moving human ...

Automatically segment the apparent contour and estimate the 2D shape of the subject's body parts (without a prior model for the human body or for the shape of the parts).

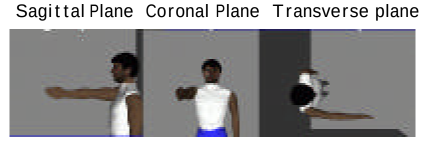
Automatically acquire a three-dimensional model of the subject's body parts.

23

Experimental Setup



- Sagittal Plane Coronal Plane Transverse plane

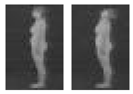


24

Human Body Model Acquisition

Protocol of movements: MovA

1. Head Motion



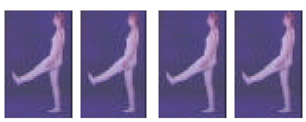
2. Left upper body extremities motions
3. Right upper body extremities motions

25

Human Body Model Acquisition

Protocol of movements: MovA (cont.)

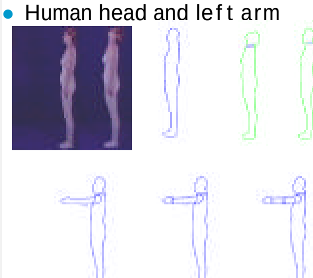
4. Left lower body extremities motions
5. Right lower body extremities motions



26

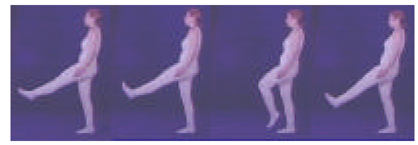
Results

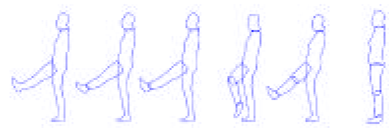
- Human head and left arm



27

Results - Human leg

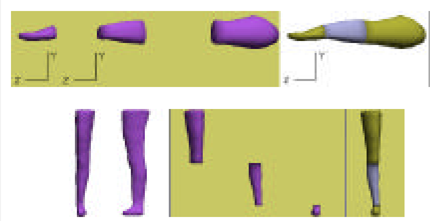




28

Results

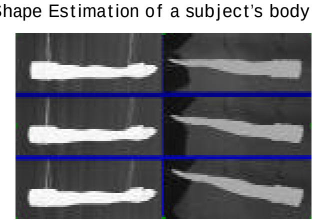
- 3D models for the arm and the leg



29

Validation and Performance Analysis

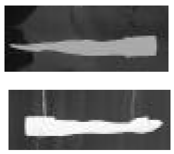
- 3D Shape Estimation of a subject's body parts

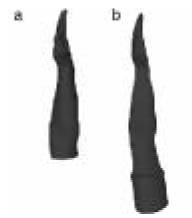


30

Validation and Performance Analysis

3D Shape Estimation of a subject's body parts





min error : 0.001 mm
max error : 3.736 mm
mean : 1.459 mm
std dev : 1.170 mm

31

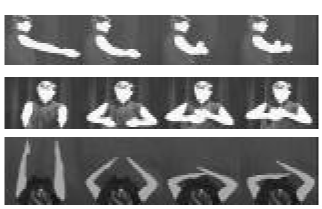
Overview

- Motivation
- Theoretical Framework
- Human Motion Analysis
  - Inferring Structure in 2D
  - Human Body Model Acquisition
  - Human Body Tracking
  - Estimating Anthropometry and Pose from a Single Camera
- Biomedical Data Analysis
- Conclusion

32

Human Motion Capture

Given image sequences of a moving human…
Estimate over time the 3D position and orientation of a subject's body parts.



33

Challenges

- Humans perform complex 3D non-rigid motions
- Body parts may not be visible from certain viewpoints

34

3D Model-based tracking

Input
- Image sequences of the moving human from three views, and
- The 3D models of the subject's body parts (as obtained with our method)

Output
- The 3D position and orientation over time of each of the subject's body parts

35

Human Motion Capture

Advantages of our approach
- Obviates the need for markers or special equipment
- Model obtained from observations
- Mitigates difficulties arising due to occlusion among body parts
- Selects a subset of the cameras in an active and time varying fashion

36

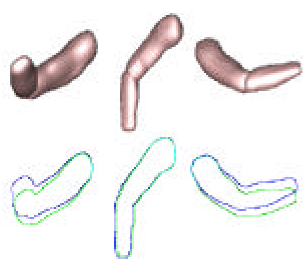
Model-Based Tracking: Steps

Steps
- Predict
- Select
- Match
- Update

37

Model-Based Tracking: Select

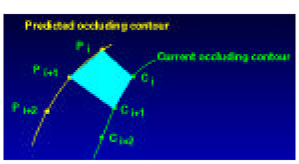
Observability Index



38

Model-Based Tracking: Select

Predicted occluding contour

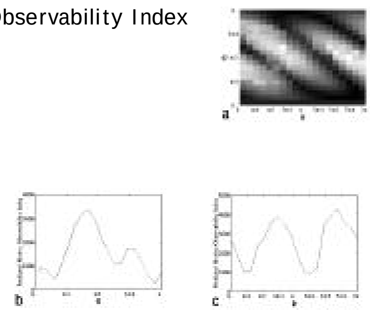


Observability Index $(I) = \sum area(c_i, c_{i+1}, P_{i+1}, P_i)$

39

Model-Based Tracking: Select

- Observability Index



40

Model-Based Tracking: Update

Lagrange equations of motion

$$\ddot{\mathbf{q}} + \dot{\mathbf{q}} = \mathbf{f}_q$$

where

$q(t)$: the generalized coordinate vector

$f_q(t)$: generalized external forces

41

Model-Based Tracking: Predict

Extended Kalman Filter

$$\begin{bmatrix}\ddot{q}\\ \dot{q}\end{bmatrix} = \begin{bmatrix}-1 & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}\dot{q}\\ q\end{bmatrix}(t) + w(t)$$

$$z(t) = h\left(\begin{bmatrix}\dot{q}\\ q\end{bmatrix}(t)\right) + v(t)$$

$z(t)$: vector of observations

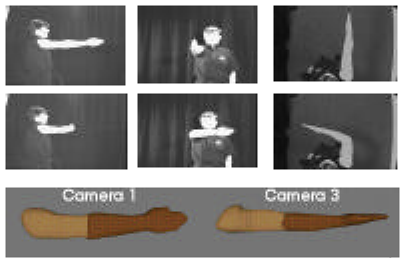
$h(t)$: nonlinear function which relates the input data to the model's state

$w(t)$: modeling error

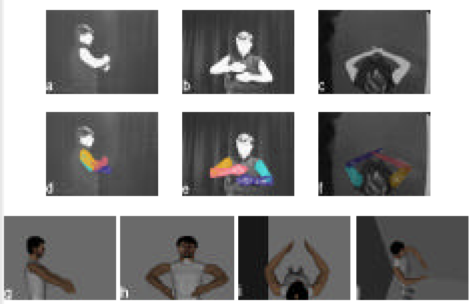
42

Human Body Model Acquisition

Frame	Camera1	Camera2	Camera3
001			
035			

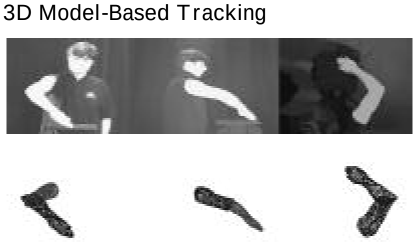


Human Motion Capture



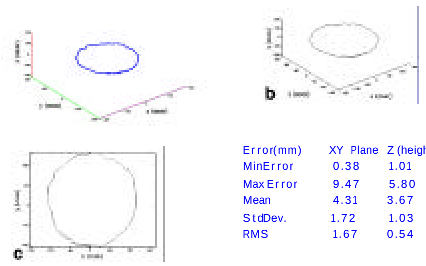
Validation and Performance Analysis

3D Model-Based Tracking



Validation and Performance Analysis

3D Model-Based Tracking

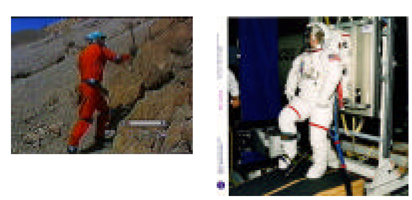


Error(mm)	XY Plane	Z (height)
MinError	0.38	1.01
Max Error	9.47	5.80
Mean	4.31	3.67
StdDev.	1.72	1.03
RMS	1.67	0.54

Video Presentation

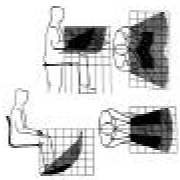
Tracking Using Monocular Images

- There are several applications for which the video recordings from only one view are available



Motivation

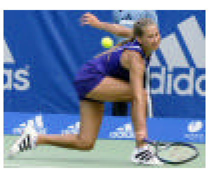
- Performance measurement for human factors engineering

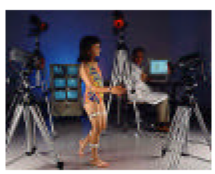


49

Motivation (Cont.)

- Posture and gait analysis for training athletes and physically challenged individuals





http://www.motionanalysis.com

50

Motivation (Cont.)

- Human body, hands and face animation

http://ligwww.epfl.ch/

51

Motivation (Cont.)

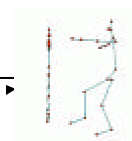
- Automatic annotation of human activities in video databases

52

Problem Statement

Given a set of points in an image that correspond to the projection of landmark points of a human subject …

estimate both the anthropometric measurements (up to a scale) of the subject and his/her pose that best match the observed image

Image →  ← 3D Human Model

(up to scale)

53

Our Approach

Novelty

✓ Using anthropometric statistics to constrain the estimation process

Advantages

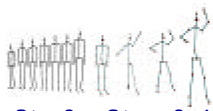
✓ Estimation of both anthropometry and pose simultaneously

✓ Able to estimate anthropometry and pose from a single image

54

Overview

Novelty: Using anthropometric statistics to constrain the estimation process



Step 1

Selection of projected landmarks on the image

Step2 Steps 3, 4 ...

- Initial Anthropometric Estimates
- Initial Pose Estimates
- Iterative Minimization over lengths and angles

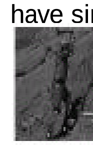
Output

Human model

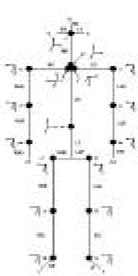
Step 1: Selection of projected landmarks

Through a simple interface, the user:

- Selects the projection of visible landmarks of the subject's body
- Marks
  - segments whose orientation is almost parallel to the image plane
  - pairs of segments that have similar orientation



Human Body Model



ID	Joint	From	To	DOF	PR
at	atlanto occipital	NK	HD	Tz*Rz*Ry*Rx	3
sp	solar plexus	UT	NK	Tz*Ry*Rz*x	2
la	left ankle	LLL	LF	Tx*Rz*Rx*Ry	4
lc	left clavicle	UT	LC	Tz*Rx*Ry	3
le	left elbow	LUA	LLA	Tz*Ry	5
lh	left hip	LT	LUL	Tz*Rz*Rx*Ry	2
lk	left knee	LUL	LLL	Tz*R-y	3
ls	left shoulder	LC	LUA	Tz*Rz*Rx*Ry	4
lw	left wrist	LLA	LHD	Tz*Ry*Rx*Rz	6
ra	right ankle	RLL	RF	Tx*R-z*R-x*Ry	4
rc	right clavicle	UT	RC	Tz*R-x*Ry	3
re	right elbow	RUA	RLA	Tz*Ry	5
rh	right hip	LT	RUL	Tz*R-z*R-x*Ry	2
rk	right knee	RUL	RLL	Tz*R-y	3
rs	right shoulder	RC	RUA	Tz*R-z*R-x*Ry	4
rw	right wrist	RLA	RHD	Tz*Ry*R-x*R-z	6
wt	waist	LT	UT	Tz*Ry*Rz*Rx	1

- 22 segments, 17 joints and 64 DoF

Output

- Image coordinates of the selected projected landmarks
- a set of ratios (of projected lengths) using the segments selected by the user

Family of Human Body Models

- 2187 human body models based on anthropometric statistics

- The cadre family is a representation of the population distribution which spans the space to capture a significant amount of the variance

Step 2: Initial Anthropometric Estimates

Input: a set of ratios using the segments selected by the user

Output: the initial human model q^* from our cadre family of 2187 human models.

$$q^* = \arg\min_{q=1,\ldots,139} \sum_{i,j\in I} \left(r_{ij}^q - p_{ij}\right)^2$$

where $r_{ij}^q = \frac{l_i^q}{l_j^q}, i<j$ are the ratios of the segments of each cadre family member that correspond to the segments selected

Steps 3 and 4: Estimates for pose and anthropometry

Goal

- ✓ Minimize the discrepancy between the synthesized appearance of the Stick Model (for that pose) and the image data of the subject in the given image

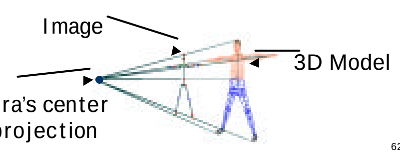
$$\min f(x_j)$$
$$L_j \le x_j \le U_j,\ j = 1,\ldots,K$$

where x_j can be an angle or a length or a ratio, and L_j and U_j are its lower and upper values.

61

The Objective Function

- The sum of squared distances between the Stick model's site and the closest point from the line formed by the camera's center of projection and its corresponding landmark



62

Minimization process

To guide the minimization process to a solution for a pose that is anthropometrically plausible, we apply:

- ✓ a geometric method for the initial pose estimation
- ✓ a hierarchical solver
- ✓ various constraints

63

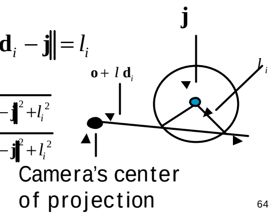
Initial Pose Estimates

We use a geometric method for providing two initial guesses for the pose of some segments as follows:

$$\|\mathbf{o} + l\,\mathbf{d}_i - \mathbf{j}\| = l_i$$

Solutions

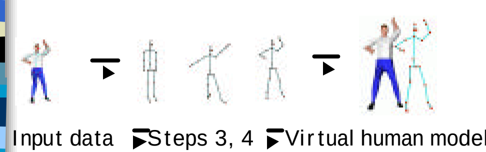
$$l_1 = \mathbf{d}_i \cdot (\mathbf{j}-\mathbf{o}) + \sqrt{[\mathbf{d}_i \cdot (\mathbf{j}-\mathbf{o})]^2 - \|\mathbf{o}-\mathbf{j}\|^2 + l_i^2}$$
$$l_1 = \mathbf{d}_i \cdot (\mathbf{j}-\mathbf{o}) - \sqrt{[\mathbf{d}_i \cdot (\mathbf{j}-\mathbf{o})]^2 - \|\mathbf{o}-\mathbf{j}\|^2 + l_i^2}$$



64

Hierarchical Solver

We assign a priority to each joint and site, and we schedule the optimization process



65

Constraints

Three classes of constraints are applied:

- Constraints derived from the joint limit information associated with the range of motion of each joint,
- Constraints that enforce the symmetry between the left and right sides of the subject, and
- Constraints that enforce proportions.

66

Results

Synthetic Experiment

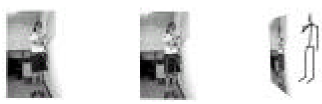
Accuracy

	$\frac{LC}{UT+LT}$	$\frac{LLA}{LUA}$	$\frac{LHP}{LUA}$	$\frac{LF}{LUL}$	$\frac{LF}{LUL}$
Actual	0.6553	0.9829	0.5700	0.6397	0.6341
Estimated	0.6517	0.9781	0.5713	0.6377	0.6329
PE %	0.5494	0.4810	0.2281	0.3090	0.1876

67

Results

Subject: Vannesa



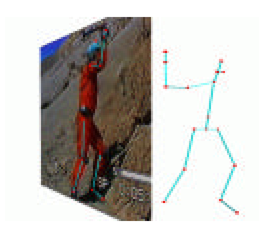
Accuracy

	$\frac{LC}{UT+LT}$	$\frac{LLA}{LUA}$	$\frac{LHP}{LUA}$	$\frac{LF}{LUL}$	$\frac{LF}{LUL}$
Actual	0.6279	0.8625	0.6949	0.5517	0.4778
Estimated	0.6266	0.8516	0.6925	0.5468	0.4767
PE %	0.1958	1.2638	0.3180	0.8957	0.2302

68

Results

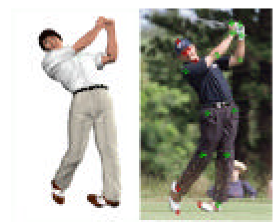
- Field work



69

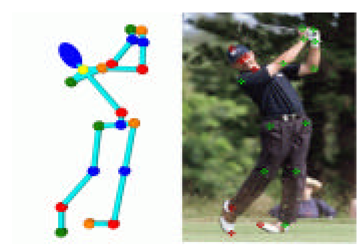
Results

- Golfer



70

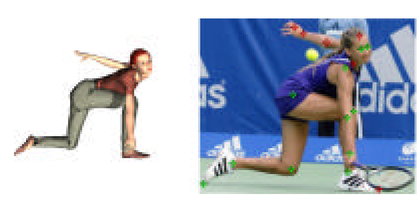
Golf



71

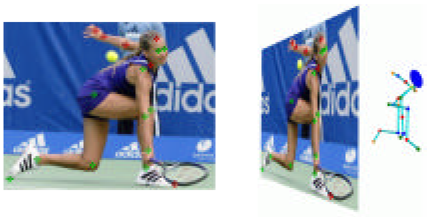
Results

- Tennis Player



72

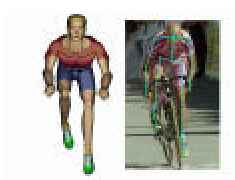
Tennis



73

Results

- Cyclist



74

VER I

Research Thrusts

- Intelligent Systems
- Computational Biomedicine
- Biomedical Robotic Systems
- Geophysical Data Analysis and Visualization

Analysis,Modeling, Simulation, Visualization
Multimodal Human Computer Interaction

75

Conclusions

- "We live in interesting times"
- Abundance of sensors
- Large volumes of information rich data
- New efficient and robust methods for analyzing, querying, visualizing and storing data

76

Our Team

- **Post-Doc**
  - Geovanni Martinez
- **Ph.D. Students**
  - Carlos Barron
  - Amol Pednekar
- **M.S. Student**
  - Anthony Do
- **Undergraduate Student**
  - Vanessa Zavaletta

77

Acknowledgements

- NSF (CAREER Award)
- NASA JSC
- Texas Higher Education Coordinating Board
- American Honda
- Shell Foundation

78